

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR LETTERS PATENT

**RESOLVING PARTIAL MEDIA TOPOLOGIES**

Inventors:

Samuel Amin

Brian D. Crites

Kirt A. Debique

Sohail Baig Mohammed

Niranjan S. Nayak

Eric H. Rudolph

Mei L. Wilson

ATTORNEY'S DOCKET NO. MS1-1852US

CLIENT'S DOCKET NO. 305164.01

**88436702960**

## **RESOLVING PARTIAL MEDIA TOPOLOGIES**

### **TECHNICAL FIELD**

[0001] The described subject matter relates to electronic computing, and more particularly to systems and methods for resolving partial media topologies.

### **BACKGROUND**

[0002] Media content has traditionally been distributed using equipment and protocols that are application-specific, and in some cases proprietary. For example, video content has traditionally been encoded in an analog format and distributed over television networks, cable networks, satellite networks, and video cassette tapes. Special purpose capture and transmission devices are required to generate the content. Similarly, special purpose receivers and display devices are required to access the content.

[0003] The widespread digitization of media (including multimedia) content, especially by the consumer segment, coupled with the growth in digital communication networks and easier methods to transfer digital content is changing the nature of media content delivery and usage. Media content can now be captured and encoded in one or more of a plurality of digital formats (e.g., MPEG, Windows Media Format, VCD, etc.) distributed over digital networks such as the internet or on digital media and accessed using general purpose computing equipment or special purpose equipment.

[0004] Digital computing devices play a central role in digital media production, encoding, distribution, and display. Microsoft Corporation of Redmond, Washington, USA, has developed a set of technologies to facilitate the use of digital media and the integration of digital media processing components (both hardware and software) with personal computers. MICROSOFT DIRECTSHOW is a digital media streaming architecture designed for digital audio, video and other types of digital data. DIRECTSHOW provides a high-level application model that enables independent hardware vendors (IHVs) and independent software vendors (ISVs) to develop streaming media applications that combine and use components from possibly different vendors and run on computers using the WINDOWS brand operating system.

[0005] Additional infrastructure to facilitate the integration of digital media components is desirable to facilitate continued development in the digital media marketplace and to increase the flexibility that users and developers have to create innovative uses of those components.

## **SUMMARY**

[0006] Implementations described herein provide systems and methods for resolving a partial media topology. In one exemplary implementation, a method for resolving a partial media topology is provided. A partial media topology that includes a plurality of nodes including at least one media source node and at least

one media sink node is received by a topology loader module. The topology loader module populates an initial upstream source list comprising the source node(s) in the partial topology. In addition, iteratively, for each source node in the source node list, the topology loader module negotiates a media type for each output of the source node with the downstream node in the partial topology, instantiates one or more intermediate nodes, connects the one or more intermediate nodes between the media source node and the media sink node, and adds the one or more intermediate nodes to the upstream source node list.

[0007] In another exemplary implementation, a system is provided. The system comprises one or more computer-readable media, and a media engine embodied on the one or more computer-readable media and configured to communicatively interact with an application to present a media presentation. The media engine is configured to use a media session to generate a partial topology, the partial topology including one or more media sources individual ones of which serving as a source of media content, and one or more media sinks configured to sink a media stream, and a topology loader to resolve the partial topology into a full topology.

[0008] In yet another implementation a system is provided. The system comprises one or more computer-readable media, and a media engine embodied on the one or more computer-readable media and configured to communicatively interact with an application to present a presentation. The media engine is

configured to use a media session to generate one or more media sources individual ones of which serving as a source of media content, and one or more media sinks configured to sink a media stream, and a topology loader to generate one or more transforms communicatively linked with one or more media sources and configured to operate on data received from the one or more media sources.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

[0009] Fig. 1 is an illustration of an environment in an exemplary implementation in which a computer provides access to a plurality of media;

[0010] Fig. 2 is a high level block diagram of a system in an exemplary implementation in which the system, implemented in software, includes an application that interacts with a media foundation to control presentation of a plurality of media;

[0011] Fig. 3 is a schematic illustration of an exemplary partial topology;

[0012] Fig. 4 is a flowchart illustrating exemplary operations for resolving a partial topology;

[0013] Fig. 5 is a flowchart illustrating exemplary operations executed for negotiating a media type with downstream nodes;

[0014] Fig. 6 is a flowchart illustrating addition exemplary operations for negotiating a media type with downstream nodes;

[0015] Fig. 7 is a flowchart illustrating operations for handling optional nodes; and

[0016] Fig. 8 is an illustration of an exemplary operating environment.

## **DETAILED DESCRIPTION**

[0017] Described herein are exemplary storage network architectures and methods for resolving partial media topologies. The methods described herein may be embodied as logic instructions on a computer-readable medium. When executed on a processor, the logic instructions cause a general purpose computing device to be programmed as a special-purpose machine that implements the described methods. The processor, when configured by the logic instructions to execute the methods recited herein, constitutes structure for performing the described methods.

### **Exemplary Environment**

[0018] **Fig. 1** is an illustration of an environment 100 in an exemplary implementation in which a computer 102 provides access to a plurality of media. The computer 102, as illustrated, may be configured as a personal computer (PC). The computer 102 may also assume a variety of other configurations, such as a mobile station, an entertainment appliance, a set-top box communicatively coupled to a display device, a wireless phone, a video game console, a personal digital

assistant (PDA), and so forth. Thus, the computer 102 may range from a full resource device with substantial memory and processor resources (e.g., PCs, television recorders equipped with hard disk) to a low-resource device with limited memory and/or processing resources (e.g., a traditional set-top box). An additional implementation of the computer 102 is described in relation to Fig. 6.

[0019] The computer 102 may obtain a variety of media from a variety of media sources. For example, the computer 102 may locally store a plurality of media 104(1), ..., 104(k), ..., 104(K). The plurality of media 104(1)-104(K) may include an assortment of audio and video content having various formats. Further, the media 104(1)-104(K) may be obtained from a variety of sources, such as from an input device, from execution of an application, and so on.

[0020] The computer 102, for instance, may include a plurality of applications 106(1), ..., 106(n), ..., 106(N). One or more of the plurality of applications 106(1)-106(N) may be executed to provide media, such as documents, spreadsheets, video, audio, and so on. Additionally, one or more of the plurality of applications 106(1)-106(N) may be configured to provide media interaction, such as encoding, editing, and/or playback of the media 104(1)-104(K).

[0021] The computer 102 may also include a plurality of input devices 108(1), ..., 108(m), ..., 108(M). One or more of the plurality of input devices 108(1)-108(M) may be configured to provide media for input to the computer 102. Input device 108(1), for instance, is illustrated as a microphone that is configured

to provide an input of audio data, such as a voice of the user, a song at a concert, and so on. The plurality of input devices 108(1)-108(M) may also be configured for interaction by a user to provide inputs that control execution of the plurality of applications 106(1)-106(N). For example, input device 108(1) may be utilized to input voice commands from the user, such as to initiate execution of a particular one of the plurality of applications 106(1)-106(N), control execution of the plurality of applications 106(1)-106(N), and so forth. In another example, input device 108(m) is illustrated as a keyboard that is configured to provide inputs to control the computer 102, such as to adjust the settings of the computer 102.

[0022] Further, the computer 102 may include a plurality of output devices 110(1), ..., 110(j), ..., 110(J). The output devices 110(1)-110(J) may be configured to render media 104(1)-104(K) for output to the user. For instance, output device 110(1) is illustrated as a speaker for rendering audio data. Output device 110(j) is illustrated as a display device, such as a television, that is configured to render audio and/or video data. Thus, one or more of the plurality of media 104(1)-104(K) may be provided by the input devices 108(1)-108(M) and stored locally by the computer 102. Although the plurality of input and output devices 108(1)-108(M), 110(1)-110(J) are illustrated separately, one or more of the input and output devices 108(1)-108(M), 110(1)-110(J) may be combined into a single device, such as a television having buttons for input, a display device, and a speaker.



[0023] The computer 102 may also be configured to communicate over a network 112 to obtain media that is available remotely over the network 112. The network 112 is illustrated as the Internet, and may include a variety of other networks, such as an intranet, a wired or wireless telephone network, a broadcast network, and other wide area networks. A remote computer 114 is communicatively coupled to the network 112 such that the remote computer 114 may provide media to the computer 102. For example, the remote computer 114 may include one or more applications and a video camera 116 that provides media, such as home movies. The remote computer 114 may also include an output device to output media, such as the display device 118 as illustrated. The media obtained by the computer 102 from the remote computer 114 over the network 112 may be stored locally with the media 104(1)-104(K). In other words, media 104(1)-104(K) may include locally stored copies of media obtained from the remote computer 114 over the network 112.

[0024] Thus, the computer 102 may obtain and store a plurality of media 104(1)-104(K) that may be provided both locally (e.g., through execution of the plurality of applications 106(1)-106(N) and/or use of the plurality of input device 108(1)-108(M)), and remotely from the remote computer 114 (e.g., through execution of application and/or use of input devices). Although the plurality of media 104(1)-104(K) has been described as stored on the computer 102, the media 104(1)-104(K) may also be provided in “real-time”. For example, audio data may

be streamed from the input device 108(1), which is illustrated as a microphone, without storing the audio data.

[0025] The computer 102 includes a timeline generator 120 that, when executed on the computer 102, generates a media timeline 122. For example, the timeline generator 120 may be configured as an application that exposes one or more software components that may be used to generate the media timeline 122, such as through a user interface by a user. As previously described, the media timeline 122 provides a technique for a user to define a presentation of stored and/or real-time media from the plurality of media sources. For example, the media timeline 122 may describe a collection of media that was obtained from the input devices 108(1)-108(M), the applications 106(1)-106(N), and/or the remote computer 114. The user may utilize one or more of the input devices 108(1)-108(M) to interact with the timeline generator 120 to define groupings and/or combinations of the media 104(1)-104(K). The user may also define an order and effects for presentation of the media 104(1)-104(K). A timeline source 124 may then be executed on the computer 102 to render the media timeline 122. The media timeline 122, when rendered, provides the expressed groupings and/or combinations of the media 104(1)-104(K) for rendering by one or more of the plurality of output devices 110(1)-110(J). Additionally, the timeline generator 120 may also programmatically generate the media timeline 122 as is described in greater detail in the following implementation.

[0026] **Fig. 2** is a high level block diagram of a system 200 in an exemplary implementation in which the system 200, implemented in software, includes an application 202 that interacts with a media foundation 204 to control presentation of a plurality of media 206(g), where “g” can be any number from one to “G”. The media foundation 204 may be included as a part of an operating system to provide playback of the media 206(g) such that applications that interact with the operating system may control playback of the media 206(g) without “knowing” the particular details of the media formats. The media 206(g) may be provided from a variety of sources, such as from the media 104(1)-104(K) of Fig. 1, through execution of the applications 106(1)-106(N), use of the input devices 108(1)-108(M), output devices 110(1)-110(J), and so on.

[0027] The application 202, which may be the same as or different from applications 106(1)-106(N) of Fig. 1, interacts with a media engine 208 to control the media 104(1)-104(K). In at least some embodiments, the media engine 208 serves as a central focal point of the application 202 that desires to somehow participate in a *presentation*. A presentation, as used in this document, refers to or describes the handling of media. In the illustrated and described embodiment, a presentation is used to describe the format of the data on which the media engine 208 is to perform an operation. Thus, a presentation can result in visually and/or audibly presenting media, such as a multimedia presentation in which both audio and accompanying video is presented to user within a window rendered on a

display device, such as output device 110(j) of Fig. 1 that is illustrated as a display device that may be associated with a desk-top PC. A presentation can also result in writing media content to a computer-readable medium such as a disk file. Thus, a presentation is not limited to scenarios in which multimedia content is rendered on a computer. In some embodiments, operations such as decoding, encoding and various transforms (such as transitions, effects and the like), can take place as a result of a presentation.

[0028] In an embodiment, the media foundation 204 exposes one or more application program interfaces that can be called by the application 202 to interact with the media 206(g). For example, the media foundation 204 may be thought of as existing at an “infrastructure” level of software that is executed on the computer 102 of Fig. 1. In other words, the media foundation 204 is a software layer used by the application 202 to interact with the media 206(g). The media foundation 204 may be utilized to control a number of aspects of the media 206(g), such as output, rendering, storage, and so on. Thus, the media foundation 204 may be utilized such that each application 202 does not have to implement separate code for each type of media 206(g) that may be used in the system 200. In this way, the media foundation 204 provides a set of reusable software components to do media specific tasks.

[0029] The media foundation 202 may utilize several components among which include the media timeline 122, the timeline source 124, a media source

210, a media processor 212, a media session 214, the media engine 208, a source resolver 216, one or more transforms 218, one or more media sinks 220, 222, and so on. One advantage of various illustrated and described embodiments is that the system 200 is a pluggable model in the sense that a variety of different kinds of components can be utilized in connection with the systems described herein. Also included as a part of system 200 is a destination 224, which is discussed in more detail below. In at least one embodiment, however, the destination 224 is an object that defines where a presentation is to be presented (e.g. a window, disk file, and the like) and what happens to the presentation. That is, the destination may correspond to one or more of the media sinks 220, 222 into which data flows.

[0030] The media timeline 122 employs a timeline object model which provides a way for a user to define a presentation based on media that is rendered by the timeline source 124. The media timeline 122 may range from a sequential list of media files to more complex forms. For example, the media timeline 122 may employ file structures, such as SMIL and AAF, to express media playback experiences that include transitions between media, effects, and so on. The application 202, for instance, may be configured as a media player that can play a list of songs, which is commonly referred to as a playlist. As another example, in an editing system a user may overlay one video over the other, clip a media, add effect to the media and so forth. Such groupings or combinations of media may be expressed using the media timeline 122.

[0031] The media source 210 is utilized to abstract a provider of media. The media source 210, for instance, may be configured to read a particular type of media from a particular source. For example, one type of media source might capture video from the outside world (a camera), and another might capture audio (a microphone). Alternately or additionally, the media source 210 may read a compressed data stream from disk and separate the data stream into its compressed video and compressed audio components. Yet another media source 210 might obtain data from the network 112 of Fig. 1. Thus, the media source 210 may be utilized to provide a consistent interface to acquire media.

[0032] The media source 210 provides one or more media presentation 226 objects (media presentation). The media presentation 226 abstracts a description of a related set of media streams. For example, the media presentation 226 may provide a paired audio and video stream for a movie. Additionally, the media presentation 226 may describe the configuration of the media source 210 at a given point in time. The media presentation 226, for instance, may contain information about the media source 210 including descriptions of the available streams of the media source 210 and their media types, e.g. audio, video, MPEG, and so on.

[0033] The media source 210 may also provide a media stream 228 object (media stream) which may represent a single stream from the media source 210 which can be accessed by the application 202, i.e. exposed to the application 202.

The media stream 228 thus allows the application 202 to retrieve samples of the media 206(g).

[0034] In the media foundation 204, therefore, the media source 210 is defined as a software component which outputs samples for a presentation. The timeline source 124 interprets the media timeline 122, but at the same time, may also act in a manner similar to the media source 210. For example, the timeline source 210 may be utilized to hide the intricacies of rendering the media timeline 122 to provide media described by the media timeline 122 from other components of the media foundation 204.

[0035] The media processor 212 manages data flow in a topology 230. The topology 230 defines how data flows through various components for a given presentation. A “full” topology includes each of the components, e.g. software modules, used to manipulate the data such that the data flows with the correct format conversions between different components. When a topology is created, the user might choose to create it partially. This partial topology is not sufficient, by itself, to provide a final presentation. Therefore, a component called the topology loader 232 may take the partial topology and convert it into a full topology by adding the appropriate data conversion transforms between the components in the partial topology.

[0036] In the topology 230, for example, data generally originates at the media source 210, flows through one or more transforms 218, and proceeds into

one or more media sinks 220, 222. Transforms 218 can include any suitable data handling components that are typically used in presentations. Such components can include those that uncompress compressed data and/or operate on data in some way, such as by imparting an effect to the data, as will be appreciated by the skilled artisan. For example, for video data, transforms can include those that affect brightness, color conversion, and resizing. For audio data, transforms can include those that affect reverberation and re-sampling. Additionally, decoding and encoding can be considered as transforms.

[0037] Media sinks 220, 222 are typically associated with a particular type of media content. Thus, audio content might have an associated audio sink such as an audio renderer. Likewise, video content might have an associated video sink such as a video renderer. Additional media sinks can send data to such things as computer-readable media, e.g. a disk file and the like.

[0038] The media session 214 is a component which may schedule multiple presentations. Therefore, the media processor 212 may be used to drive a given presentation, and the media session 214 utilized to schedule multiple presentations. The media session 214, for instance, may change topologies that are rendered by the media processor 212. For example, the media session 214 may change from a first topology that is rendered on the media processor 212 to a second topology such that there is no gap between the renderings of samples from the consecutive presentations that are described by the respective topologies. Thus, the media



session 214 may provide a seamless user experience as the playback of the media moves from one presentation to another.

[0039] The source resolver 216 component may be utilized to create a media source 210 from URLs and/or byte stream objects. The source resolver 216 may provide both synchronous and asynchronous ways of creating the media source 210 without requiring prior knowledge about the form of data product by the specified resource.

[0040] In at least one embodiment, the media foundation 204 is utilized to abstract away the specific details of the existence of and interactions between various components of the media foundation 204. That is, in some embodiments, the components that are seen to reside inside the media foundation 204 are not visible, in a programmatic sense, to the application 202. This permits the media foundation 202 to execute so-called “black box” sessions. For example, the media engine 208 can interact with the media session 214 by providing the media session certain data, such as information associated with the media (e.g. a URL) and the destination 224, and can forward the application’s 202 commands (e.g. open, start, stop and the like) to the media session 214. The media session 214 then takes the provided information and creates an appropriate presentation using the appropriate destination.

[0041] The media foundation 204 may also include a timeline plugin 234. The timeline plugin 234 may be utilized such that different media timeline file

formats may be “plugged-in” to the media foundation 204. For example, a bytestream plugin 236 may be written for a format in question and registered with the media foundation 204. The source resolver 216 may then invoke a bytestream plugin 236 when a file of that type is opened. In turn the bytestream plugin 236 can parse the file, create a media timeline 122 representing the presentation described in the file, and create a timeline source 124 for it. In general, the bytestream plugin 236 is responsible for reading the raw bytestream and creating a media source 208 for it. In an implementation, the remaining components of media foundation 204 are not made aware that the media source created in this instance is a timeline source 124. Therefore, the timeline source 124 is treated like any other media source 208. In an implementation, a bytestream plugin 236 that can parse a media timeline 122 and create a timeline source 124 is referred to as a timeline plugin.

[0042] The timeline plugin 234 may also provide an interface such that the application 202 may interact with the timeline plugin directly, such as to load and save the media timeline 122 from or to a file. For example, the timeline plugin 234 may be created and then called to initiate a load function to provide a bytestream. The timeline plugin 234 may then parse the file and create a root node and any additional nodes to create the media timeline 122. The timeline plugin 234 may also be used to persist the media timeline 122 to different formats. For example, the application 202 may create the media timeline 122 programmatically.

In other words, the application may act as the timeline generator 120 of Fig. 1. The application 202 may then create a timeline plugin for ASX files, and ask the timeline plugin to save the media timeline 122 in the ASX format. In another example, a user can open an m3u file, i.e. a playlist file format for specifying multiple MP3 files, get the media timeline 122 from it, and then ask the timeline plugin to save the media timeline 122 in the ASX format. In this way, the media foundation 204 may expose a plurality of software components that provide media functionality over an application programming interface for use by the application 202.

[0043] Given the description of the system of Fig. 2, the discussion that follows provides a general overview of a typical multimedia scenario, along with a description of the roles that the media engine 208 and media session 214 plays in driving the presentation. In the discussion that follows, each of the media engine (and its role) and media session (and its role) are discussed in sections under their own respective headings—*i.e.*, “Media Engine Work” and “Media Session Work”.

### **Media Engine Work**

[0044] In accordance with one embodiment, the work that the media engine 208 performs during a presentation can be categorized, generally, under a number of different headings which appear below. The categories of media engine work include source resolution, setting up the media session, partial topology resolution,

topology resolution and activation, presentation control, new presentations, and output changes.

### Source Resolution

[0045] Source resolution pertains to the process by which the media engine 208 causes the appropriate media source to be created for the particular type of data that is to be read and subsequently processed by the system. Thus, this process obtains a media source from which the multimedia data can be read. This process is relevant when, for example, the OpenURL or OpenByteStream methods (discussed above and below) are called to open the multimedia. In either case, the media engine 208 passes the URL or the Byte Stream, respectively, to a component known as a source resolver. If the source resolver is given a URL, then it looks at the scheme of the URL (*e.g.*, file://, http://, etc) to create a Byte Stream that will read from the specified location.

[0046] In both cases, the source resolver is able to look at the contents of the Byte Stream to determine the format of the bits (*e.g.*, ASF, AVI, MPEG, etc) so that a media source can be instantiated that will understand that format. The other Open functions discussed above and below specify the media source directly.

### Setting up the Media Session

[0047] During this process, the media engine asks the media source that is created for a presentation descriptor. In some embodiments, the presentation descriptor may specify that a custom media session is to be used. In many cases, however, custom media sessions may not be used in which case a default media session can be instantiated.

### Partial Topology Creation

[0048] During partial topology creation, the media engine 208 obtains a presentation descriptor from the media source(s) 210 and notifies the application 202 of that particular presentation via the event `MENewPresentation`. If the application is interested in using that event to configure a destination, the media engine 208 waits for the application to finish handling the event.

[0049] The media engine 208 then negotiates with the application-provided destination and the destination can create one or more media sinks for the outputs of the presentation. In some embodiments, media sink(s) 220, 222 can have already been created and the destination simply hands them over to the media engine.

[0050] The media engine 208 invokes the media processor 212 to construct a “partial topology” that the media engine indicates the source media streams and the output stream sinks, without necessarily specifying the transforms that will be

needed to get there. Thus, referring to the Fig. 2 illustration, at this point in the process, the media engine 208 has created a partial topology, a media processor 212 and a media session. Media engine has instantiated or referenced one or more media source(s) 210 and media sink(s) 220, 222.

[0051] An exemplary partial topology is depicted in **Fig. 3A**. Referring to Fig. 3, a partial topology 300 specifies an audio source node 310 and an audio render node 320, a video source node 330 and a video render node 340.

#### Full Topology Resolution

[0052] In performing the topology resolution process, the media session 214 can invoke a component referred to herein as a topology loader 232. The topology loader 232 implements logic instructions to determine which transforms 218 are necessary or desirable to provide the data from the media source(s) 210 to the media sink(s) 220, 222.

[0053] Transforms 218 can comprise any suitable data handling components that are typically used in presentations. Such components can include those that uncompress compressed data, and/or compressed uncompressed data, and/or operate on data in some way, such as by imparting an effect to the data, as will be appreciated by the skilled artisan. For example, for video data, transforms can include those that affect brightness, color conversion, and resizing. For audio data,

transforms can include those that affect reverberation and resampling. Additionally, decoding and encoding can be considered as transforms.

[0054] An exemplary full topology is depicted in **Fig. 3B**. Referring to Fig. 3B, a full topology specifies an audio source node 310, an audio decoder 312, a resampler 314 and an audio render node 320. The full topology further comprises a video source node 330, a video decoder 332, a color converter 334, and a video render node 340.

[0055] Topology loader 232 and its operations are described in detail herein.

#### Topology Resolution and Activation

[0056] In accordance with one embodiment, during topology resolution and activation, the media engine 208 asks the media session 214 to resolve the partial topology into a fully specified topology. The media engine 208 then sets the new fully-specified topology on the media session 230, which gives it to the media processor 212. As an example, consider that the media source that is created is one that reads a compressed WMV file. The sinks, on the other hand, are not configured to handle compressed data. Thus, during topology resolution, the media session ascertains which transforms are necessary to provide the compressed data from the WMV file to the sinks and creates the appropriate transforms which, in this case, might comprise a decompressor and possibly resizers, color converters, resamplers, and the like.

[0057] In another embodiment, resolution and activation can be combined into a single operation. Specifically, the media engine 208 can set a partial topology on the media session 214 and the media session itself can resolve the partial topology into a fully-specified topology which it then provides to the media processor 212.

#### Media Processor Creation

[0058] The media session 214 is responsible for creating the media processor 212. That is, the media session 214 owns the media processor 212. When the topology is set on the media session 230, the media session 230, in turn, sets the topology on the media processor 212. The media processor 212 follows the data flow laid out by the topology to transform data from the media source(s) 210 to the particular formats that are needed by the media sink(s) 220, 222.

#### Time Source Selection

[0059] One of the functions that the media session 214 can perform pertains to time source selection. Specifically, upon starting a presentation, the media session 214 can make a determination as to which of the available time sources will be used to drive the presentation. Each component can then run its part of the presentation in synchronization with the time from the time source ascertained by the media session. The time source is also used in the presentation clock (owned



by the media engine but given to the media session) for the purposes of reporting progress of the presentation.

[0060] Media sinks, such as sinks 220, 222 may optionally offer a time source. Typically, the audio renderer (*i.e.*, audio sink) can offer a time source, and the time on the time source will be dictated by the audio device on the particular machine on which the presentation is presented. It is to be appreciated, however, that other media sinks may do so as well. In addition, a particular media source, *e.g.*, live media sources such as device capture and network sources, may also provide some concept of time. In one embodiment, the media session takes care of attempting to make the time source it chooses run at a similar rate to that of the live media source. In one embodiment, the media session 214 can decide which of the time sources is the “highest priority” time source, and this time source is used by the main presentation clock, to which all clock-aware components synchronize their presentations.

#### Presentation Control

[0061] As noted above, the media session 214 can receive method calls to Start, Stop, and Pause from the media engine 220. These calls typically correspond to the applications calls that are made on the media engine 220.

[0062] The media session 214 can control the presentation via a Presentation Clock that it receives from the media engine 220. Starting, stopping

and/or pausing the Presentation Clock results in all media sink(s) 220, 222 receiving notifications thereof and reacting appropriately. The media session 214 starts, stops, and/or pauses the media processor 212 by respectively calling its start, stop, and/or pause methods directly.

[0063] The media session 214 is configured, in this embodiment, to send an event to the media engine 208 after a given operation has been completed by all streams.

#### New Presentations and Output Changes

[0064] In accordance with this embodiment, media session 214 is responsible for forwarding media processor's 240 notification of an upcoming new presentation to media engine 208 and participating with topology resolution and activation, as described above in connection with the media engine.

#### Time Line Processing

[0065] In accordance with one embodiment, media session 214 is configured to reduce glitches at presentation startup time and when transitioning between presentations in a timeline.

[0066] In accordance with this embodiment, at startup time, media session 214 will get the first few samples of media data from media processor 212 and deliver them to the media sinks 220, 222 before starting the clock associated with

the presentation. This process uses a special “prerolling” capability on the media sinks that allows the media sinks to receive data before actually being started. In this embodiment, it is only after the media sinks receive data via the pre-rolling capability that media session 214 will start the presentation clock.

[0067] Because the media sinks 220, 222 have already received the initial data of the data stream, the chances that the media sinks will fall behind (i.e. referred to as a “glitch”) at the beginning of the presentation are greatly reduced if not eliminated all together. This can effectively provide for a generally seamless presentation start.

[0068] At presentation transition boundaries (i.e. when changing from one presentation to another), media session 214 is configured to attempt to make the transition seamless, i.e. without interruption between the end of the first presentation and the beginning of the second. In accordance with this embodiment, the media session 214 accomplishes this by applying some logic to ensure that the “seamless stream” plays continuously throughout the transition, without waiting for other streams in the presentation to complete (which may cause a glitch during the transition).

### Content Protection

[0069] In accordance with one embodiment, system 200 and more generally, systems that employ a media session component as described in this document, can

employ techniques to ensure that media content that is the subject of a presentation is protected in accordance with rights that may be associated with the content. This concept is also referred to by some as “digital rights management”.

[0070] Specifically, certain multimedia content may have specific rights associated with it. For example, the content provider may wish to restrict playback of this content to the use of only known, trusted transforms, media sinks and other components. Accordingly, content protection information associated with the media content may, but need not then be embedded in the content as will be appreciated by the skilled artisan. In accordance with this embodiment, media session 214 is configured to respect any content protection requirements by validating all of the components that are being inserted into the pipeline and by making sure that the components are allowed and will be performing allowed actions on the content. Validation can take place by any suitable measures. For example, in validating the component, the media session can then validate the component’s signature, and that the signing authority is a trusted authority.

[0071] In accordance with one embodiment, the media session 214 can create a protected media path for such content. The protected media path is configured to make it very difficult if not impossible for unauthorized third parties to intercept the data flowing through the pipeline.

### Media Engine Configuration

[0072] One of the more common scenarios in which the above-described systems and methods can be employed pertains to setting up a simple playback of a multimedia presentation. From the application's point of view, it is desirable for the application to be able to accomplish the following steps in order to configure a multimedia presentation. The application should be able to create a media engine and a playback or presentation destination. The application should also be able to provide a handle to the presentation destination, *e.g.*, a window in which a video for the presentation should be rendered. The application should also be able to call `IMFMediaEngine::OpenURL`, to supply a URL to the multimedia file to be presented, as well as a pointer to the playback destination. With these capabilities, the application can now cause the media presentation to be played back by using the `IMFMediaEngine::Start/Stop/Pause` APIs. In one embodiment, the application does not need to wait for any events to arrive as handing of these events are optional. In another embodiment, the application does handle events from the media engine for the open operation to complete.

### Exemplary Topology Loader

[0073] In an exemplary implementation topology loader 232 implements methods for converting a partial topology generated by the media processor 212 into a full topology. As used herein, the term "full topology" refers to a topology

in which all the requisite intermediate transforms are in the topology, all the input and output media types have been set on every object in the topology, and all the source nodes in the topology are ready to run. When the topology loader 232 is finished, the full topology may be processed by the media processor 212.

[0074] In an exemplary implementation the topology loader 232 is a public object, in that it is intended that it will be used by an end-user to find fully specified topologies. By way of overview, a client invokes the topology loader 232 by calling a load method and providing a partial topology. The topology loader 232 enumerates all the source nodes in the partial topology and places these nodes into a queue. This queue is then processed node by node to connect each node to its outputs. The internal nodes in the partial topology are added to the queue if and only if all their input connections have been resolved. This ensures that the topology loader 232 does not try to configure the outputs of a component before configuring all of its inputs.

[0075] For every partial connection the topology loader 232 is trying to connect, the outputs of a node are connected to the inputs of a downstream node. If one or more of the nodes are compressed, then intermediate nodes to decompress the stream may be inserted.

[0076] Features and operations of an exemplary implementation of a topology loader 232 are described herein.

## External Interfaces and Methods

[0077] In an exemplary implementation topology loader 232 provides a plurality of interfaces that may be called by an external process such as, *e.g.*, the media session 214 to invoke functions of the topology loader. The following description of exemplary interfaces and accompanying methods are provided by way of disclosure.

### IMFTopoLoader Interface

```
interface IMFTopoLoader : IUnknown
{
    HRESULT SetTopologyCallback(IMFTopoologyConnectionCallback*
pTopoCallback, DWORD dwFlags );
    HRESULT SetPreferredSampleDuration(          LONGLONG
llSampleDuration );
    HRESULT Load( IMFTopoology * pInputTopo, IMFTopoology **
ppOutputTopo );
};
```

[0078] SetTopologyCallback( IMFTopoologyConnectionCallback\*  
pTopoCallback, DWORD dwFlags );

[0079] This method permits an application 202 to specify a “smart” connector callback to the topology loader 232. This “smart” connector gives the application the flexibility to influence the topology loader 232 during the process of constructing a topology. By way of example, the application may be given the chance to connect two nodes before the topology loader 232 attempts to connect the nodes. Alternatively the application may have the ability chance to create a certain DMO before the topology loader 232 attempts to create the DMO.

[0080] Load( IMFTopology \* pIn, IMFTopology \*\* ppOut , IMFTopology\* pCurrentTopo)

[0081] Given an input partial or loaded topology, this method turns converts the partial topology into a fully loaded topology. This method locates the intermediate transforms needed to provide a fully specified pipeline of sources, transforms, and sinks, and sets all the input and output media types on all the objects in the topology. If this method returns successfully, the output topology is ready for processing by the media processor or another processor.

[0082] In an exemplary implementation the third parameter pCurrentTopo can be NULL or a pointer to the preceding topology. If pCurrentTopo is specified, then it will be used for object caching, which is described in greater detail below. The objects in the output topology may be the same objects in the input topology. If the Load method completes successfully, then the input topology may be discarded.

#### IMFSampleDurationSetter Interface

```
interface IMFSampleDurationSetter : IUnknown
{
    HRESULT SetPreferredSampleDuration(LONGLONG llSampleDuration );
};
```

[0083] SetPreferredSampleDuration( LONGLONG llSampleDuration);

[0084] This method may be used by an application to specify preferred sample duration. Smaller sample duration may be specified to provide lower



latency, while bigger sample duration may be specified to provide better quality. In an exemplary implementation `llSampleDuration` is in 100ns data units.

CallBack Interface: `IMFTopologyConnectionCallback`:

[0085] In an exemplary implementation topology loader 232 comprises a callback interface that permits an application such as application 202 to influence the topology loader 232 during the process of resolving a topology. Exemplary scenarios in which this callback feature may be useful include:

[0086] Codecs: An Application could have knowledge of certain codecs which it does not want to be used or is known as buggy. The callback feature offers a way for the application to reject a codec. An application may also have knowledge of certain codecs which it prefers to use. The callback feature offers a way for the application to specify preferred codecs.

[0087] User Specified DMOs: For transform nodes, the application 202 is allowed to specify a guid on the topology node in the partial topology. Since the application's DMO may potentially be a not a real COM object (*i.e.*, cannot be instantiated through `CoCreateInstance`) or might require some setup and initialization on creation, the callback should offer the application an opportunity to instantiate its DMO itself.

[0088] Topology Loader Override: In some instances an application may need to completely override the topology loader 232. For example, in some cases,

there may be domain specific negotiations between two components of which the topology loader 232 has no knowledge. The callback feature permits an application to override the topology loader 232 in a connection and allows the application to resolve the partial topology.

[0089] **Compatibility Problems:** Certain components are known to have problems with certain media types even though they accept them. By way of example, some video renderers may accept a media type but have problems rendering it. The callback feature permits an application to have some control on the media types used with the components it uses.

#### NotifyDMOCreation Method

[0090] HRESULT NotifyDMOCreation ( const CLSID\* pGuidDMO,  
TOPOID NodeID, IUnknown \*\* pDMOUnk );

[0091] A partial topology received by the topology loader may have transform nodes which do not have the objects instantiated. In this case the nodes contain the guid of the required DMOs and the topology loader creates it using CoCreateInstance. Since a user-specified transform may not be a COM object, or the application may want to set some setting to the DMO before it is used, the topology loader will first call the topology callback method if it exists with this notification.

#### [0092] Return Codes:

S\_OK: The application has created the DMO and returned it in pDMOUnk. The topology loader will attach this DMO to the node and continue.

S\_FALSE: Topology loader will create the DMO itself the normal way.

Any other Error Code: An unexpected error has occurred. The topology loader will abort with this error.

#### NotifyConnectingNodes Method

```
[0093] HRESULT NotifyConnectingNodes( IMFTopology* pTopology,
IMFTopologyNode* pUpNode, long lOutIndex, IMFTopologyNode*
pDownNode, long lInIndex );
```

[0094] The topology loader may call this method whenever it attempts to resolve the connection between two nodes in the partial topology. The connection may require any combination of decoders, encoders, color converters, etc. to be added in between.

[0095] This method may give the application an opportunity to do the work on one of the connections (instead of the topology loader) to enforce the applications specific requirements. The application has two possible approaches: it can either set the preference types it wants on one or both sides of the connection, or it can make the connection itself.

[0096] In the first approach, the application only needs to set the preference type on the side of the connection it cares about. The application has to guarantee that the type it sets is acceptable to that component. The topology loader will only use that type when working with this component next. The application will have to realize that some components might be marked unconfigurable if they are components repeated from a previous topology that is currently running. This means that the type on the object should not be changed.

[0097] After the application sets the types, it returns an HRESULT of S\_TOPO\_TYPES\_SET, (or another code that indicates that it only set the preferred types).

[0098] In the second approach the application may implement the following steps: First, the application will need to discover the type of objects it is connecting by getting the Node type from each of the nodes. The application may need to connect each possible object according to its interface. The application may connect sources, DMOs, splitters, multiplexers, tees, or sinks. Typically the application does not need to understand all of these since in most cases the application probably only cares about a specific connection it wants to handle. For all others it can return E\_NOTIMPL indicating that it does not handle connecting this connection.

[0099] Second, the application needs to negotiate the media types between the two nodes. The application should take care not to change the type of any node marked as unconfigurable. The application may decide whether any intermediate nodes are necessary.

[00100] Third, the application may Insert any intermediate components it decides are necessary to complete this connection by creating new nodes for them, adding these nodes to the topology and setting the components and their guides to the nodes.

[00101] Fourth, the application should connect all nodes together and configure all components media types.

[00102] Fifth, the application should Return S\_OK to indicate that it has completed work on this connection and the topology loader can pass on to the next one.

[00103] If the application fails to resolve the connection and wants the topology loader to attempt this, it should clean up all nodes it added to the topology and all connections it made and then return E\_NOTIMPL.

[00104] Return Values:

S\_OK: Indicates that the application has done all the work. The Topology loader will move on to next connection to resolve

MF\_S\_TOPO\_TYPES\_SET: Indicates that the application has set preferred types on the nodes, and wants the topology loader to continue resolving this connection

E\_NOTIMPL: Indicates that the application is not interested in this connection. The topology loader continues normally

Any other error code: Indicates that the application returned an unexpected failure. The topology loader will fail returning this error.

#### NotifyCodecCreation Method

HRESULT NotifyCodecCreation (CLSID\* pGuidDMO, CLSID\* pDMOCategory, TOPOID NodeID, BOOL bLastChance);

[00105] Parameters:

pGuidDMO: Guid of Codec about to be created.

pDMOCategory: One of the following values

DMOCATEGORY\_VIDEO\_ENCODER

DMOCATEGORY\_AUDIO\_ENCODER

DMOCATEGORY\_VIDEO\_DECODER

DMOCATEGORY\_AUDIO\_DECODER

NodeID: ID of upstream node whose output is being decoded or downstream node in the case of encoding.

bLastChance: Set to true when this is the last codec available.

[00106] Return Values:

S\_OK: Indicates that the application has no problem with this codec.

MF\_E\_TOPO\_CODEC\_REJECTED: Application rejects this codec.

Any Other Error: Indicates an unexpected application problem.

NotifyMediaTypeConsidered Method

HRESULT NotifyMediaTypeConsidered (IMFMediaType\* pMediatype, IMFTopologyNode\* pNode, long lIndex, bool bOutput);

[00107] Parameters:

pMediaType: MediaType being considered by topology loader

pNode: Node Containing component from which the topology loader received this type.

lIndex: Index of Node's connection for which the topology loader got this type

bOutput: True if this is the node's output connection, false if input

[00108] Return Values:

S\_OK: Indicates that the media type is accepted.

MF\_E\_TOPO\_MEDIATYPE\_REJECTED: Indicates that the application rejects this media type.

MF\_S\_TOPO\_MEDIATYPE\_IGNORE: Indicates that the application does not care about this particular connection.

[00109] In an exemplary implementation, this notification is triggered in the following circumstances:

[00110] A) When connecting two uncompressed nodes: It would be called for every media type received from the upstream node, and every media type received from the downstream node. The application could restrict the working set and the topology loader may use the remaining types to determine which media type to apply. Thus the notification does not mean the system is definitely going to

use this media type, but that it may, so this is the application's opportunity to reject it.

[00111] B) When connecting uncompressed to uncompressed with no input types (*e.g.*, a Video Renderer): This method is called with every media type imposed on the downstream node

[00112] C) When Connecting Compressed to Compressed: This method is called for every media type tried between the upstream and downstream nodes, *e.g.*, when connecting them directly without decoding and re-encoding.

[00113] This method will not notify the application of media types between sources and decoders, encoders and sinks, or internal inserted components such as the media types the topology loader uses for the color converter and resizer *etc.*

#### NotifyReceivedPartialTopo Method

HRESULT NotifyReceivedPartialTopo( IMFTopology\* pPartialTopo );

[00114] Parameters:

pPartialTopo: a pointer to the new partial topology received by the topology loader for resolution

[00115] Return Values:

S\_OK: Indicates that the application has made the changes it needs to the partial topology, and the Topology loader may proceed with resolving the topology. It is possible that the application could have left the topology unchanged.

AnyErrorCode: Indicates that an unexpected error occurred.

[00116] This notification is triggered whenever the Load method is called on the topology loader and gives the application access to the partial topology right

before it is resolved. The application can add or remove nodes or configure components already present in the topology. The application may also need this just to gather some information from the topology. It is the application's responsibility not to change the partial topology in an incorrect way causing the resolution to fail.

### IMFTopoServices Interface

#### Overview

[00117] This interface is implemented by the media session object 230. The media session 214 is the object which creates the topology loader 232. The topology loader's lifetime is the same as that of the media session 214 so any settings set to topology loader 232 though this interface will remain active though the life of the media session 230. This interface provides topology services such as setting the callback interface to the topology loader and other topology-related services which require knowledge of the current state of the media processor 214 (*e.g.*, what is the current topology).

[00118] The IMFTopoServices interface is accessed by the application through an IMFGetService interface implemented on the media engine object 220. The application will get the IMFGetService by QIing the media engine 220. Then application calls IMFGetService::GetService with the guid MF\_TOPO\_SERVICE and IID\_IMFTopoServices. The call is forwarded by the media engine 208 to the



media session 214 which returns a pointer to this interface. The form of the IMFTopoServices Interface is as follows:

```
Interface IMFTopoServices: IUnknown
{
    HRESULT SetTopologyCallback( IMFTopologyConnectionCallback*
                                pTopoCallback, DWORD dwFlags );

    HRESULT GetCurrentFullTopology( IMFTopology** ppFullTopo );

    HRESULT SetFullTopology ( IMFTopology* pFullTopo, BOOL
                              bNeedsResolution );
}
```

#### Method Descriptions

##### SetTopologyCallback Method

```
HRESULT SetTopologyCallback( IMFTopologyConnectionCallback*
                             pTopoCallback, DWORD dwFlags );
```

[00119] This method is the same as the one defined for IMFTopoloader above. It forwards the call to the topology loader 232.

##### GetCurrentFullTopology Method

```
HRESULT GetCurrentFullTopology( IMFTopology** ppFullTopo );
```

[00120] Parameters:

[00121] ppFullTopo: Returns the current full topology.

[00122] This method will return a pointer to a clone of the current full topology which is running in the media processor 230. Since the cloned nodes will still have reference to the actual objects, the application has the responsibility not

to make changes, such as changing the types on the components, as these changes might disrupt the media processor's operation.

[00123] The application may use this to either gather information about the current topology, or make changes to the topology and reset it using the next method.

#### SetFullTopology Method

```
HRESULT SetFullTopology ( IMFTopology* pFullTopo, BOOL  
    bNeedsResolution );
```

[00124] Parameters:

[00125] pFullTopo: New Topology to be set to the media processor. Usually this will be a topology obtained through GetCurrentFullTopology and modified.

[00126] bNeedsResolution: TRUE if the application has modified the topology in a way which requires it to be re-resolved in the topology loader before being set to the media processor. If it is set to false, the topology will be set as is to the media proc.

#### Internal Interfaces and Methods

[00127] In an exemplary implementation topology loader 232 provides internal interfaces and methods used to convert partial topologies to full topologies. Exemplary interfaces and accompanying methods are as follows:

## Connecting Nodes

[00128] `ConnectNode( IMFTopologyNode* pNode )`

[00129] The Load method initially adds all the source nodes to a FIFO queue, then calls the ConnectNode interface for each of the nodes in the queue until the queue is empty.

[00130] The ConnectNode interface handles every output of a node individually. The ConnectNodes method is called with the current node and the node that should be connected to the output of the current node. If a reference to the previous full topology exists, then ReconnectNodes is called instead of ConnectNodes to attempt to use preconfigured components from the old topology.

### ConnectNodes Method

`ConnectNodes( IMFTopologyNode* pUpNode, long lOutIndex,  
IMFTopologyNode* pDownNode, long lInIndex)`

[00131] This method connects one output on an upstream node to one input on the downstream node. This method implements logic to determine whether the upstream and downstream nodes are compressed, and inserts intermediate components as necessary to deal with compression. Once a connection is established with the downstream node, the downstream node is examined to determine if all its inputs are connected. If so, then the downstream node is added to the end of the FIFO queue. Operations implemented by the ConnectNodes method are described in greater detail in connection with Fig. 5.

### ConnectUnCompToUnComp Method

[00132] This method gets all output types of an upstream node and stores them in an array. Video and audio types without a MFVideoFormat or MFAudioFormat structure attached for Video and Audio types are rejected.

[00133] This method also gets all input types of downstream node and stores them in array. Video and audio types without an MFVideoFormat or MFAudioFormat structure attached for Video and Audio types are rejected.

[00134] If there are no input types on the downstream node, then all output types are tried on the downstream node until it accepts one. If the downstream node rejects all output types and is video, then connect to color converter and try all output types of color converter on downstream node. This is the path taken when connecting to a Video Renderer.

[00135] If both the upstream node and downstream node have specified types, then an attempt is made to match the types.

### GoFromUnCompToComp / GoFromCompToUnComp Methods

[00136] These two methods are similar, except that one is on the decoding side, and the other on the encoding side. These methods use a MF Transform Enumerator to find the Decoder/Encoder needed according to the type that needs to be decoded or encoded to. If no MF Transform is found, then it attempts to use the acm/icm codecs.

[00137] The nodes are then connected, and in the case of encoding the input type of the encoder is determined and set on that input, to be used in ConnectUnCompToUnComp

#### ReconnectNodes Method

[00138] This method is called instead of ConnectNodes to connect two nodes with the help of previously created components that are in the previous old topology. If it fails, then ConnectNodes is called, and the code path proceeds normally.

#### Exemplary Operations

[00139] Figs. 4-7 are flowcharts illustrating exemplary operations implemented by topology loader 232 for resolving a partial topology into a full topology. **Fig. 4** is a flowchart illustrating exemplary operations for resolving a partial topology. Referring to Fig. 4, at operation 410 the topology loader receives a partial topology, *e.g.*, as a parameter in a Load call. At operation 415 the topology loader 232 populates working FIFO queue of topology nodes with all the source nodes in the partial topology. In an exemplary implementation this involves enumerating each of the source nodes in the partial topology and adding the source nodes to a queue for processing.

[00140] The topology loader 232 then processes the working FIFO queue sequentially, connecting each node to its outputs. Operations 420-445 describe an

iterative process for resolving a partial topology. At operation 420 the topology loader 232 determines whether the working FIFO queue is empty. If not, then control passes to operation 430, and the topology loader pops the top node off the working FIFO queue. At operation 435 the topology loader 232 negotiates the media type of the selected node's outputs with the inputs of the downstream node. This operation is explained in greater detail in Fig. 5. At operation 440 the topology loader 232 checks the downstream node to determine whether it has additional inputs. If it does, then continued processing is required and control passes back to operation 420, and the top node on the working FIFO queue is popped for further processing. If the downstream node has no more inputs, then control passes to operation 445, and the downstream node is added to the working FIFO queue for processing. Control then passes back to operation 420 and, the next node on the working FIFO queue is popped for processing.

[00141] Operations 420 through 445 continue are repeated, sequentially connecting upstream nodes to downstream nodes, until there are no nodes remaining on the working FIFO queue for processing. At this point control passes to operation 425 and the topology loader returns the fully-resolved topology to the calling entity, e.g., the media processor 230.

[00142] **Fig. 5** is a flowchart illustrating exemplary operations 500 executed for negotiating a media type with downstream nodes (operation 435). Referring to Fig. 5, at operation 510 the topology loader 232 discovers the media types of the

upstream node and the downstream node. The respective media types may be retrieved from a data structure associated with the nodes. If, at operation 515, the media types of both nodes are uncompressed, then control passes to operation 520 and the topology loader 232 invokes the ConnectUncomptoUncomp method to connect the nodes.

[00143] By contrast, if both media types are not uncompressed, then control passes to operation 525. If the upstream node is compressed and the downstream node is uncompressed, then control passes to operation 530, and the topology loader inserts a decoder to uncompress the output of the upstream node. In an exemplary implementation the topology loader 232 may invoke the GoFromCompToUnComp method to locate a decoder, which may be connected to the upstream node using the procedure outlined in Fig. 4. Control then passes to operation 520, and the topology loader may invoke the ConnectUncomptoUncomp method to connect the decoder to the uncompressed downstream node.

[00144] If, at operation 535, the upstream node is uncompressed and the downstream node is compressed, then control passes to operation 540, and the topology loader 232 inserts a encoder to compress the output of the upstream node. In an exemplary implementation the topology loader 232 may invoke the GoFromUnCompToComp method to locate a encoder, which may be connected to the downstream node using the procedure outlined in Fig. 4. Control then passes

to operation 520, and the topology loader may invoke the ConnectUncomptoUncomp method to connect the upstream node to the encoder.

[00145] If test applied at operation 535 fails, then both the upstream and the downstream nodes are compressed media types. In this event the topology loader connects a decoder to the output of the upstream node to decompress the upstream node output, e.g., by invoking the GoFromCompToUnComp method. In addition, the topology loader 232 connects and encoder to the input of the downstream node, e.g., by invoking the GoFromUnCompToComp method. The two transforms may then be connected, e.g., by invoking the ConnectUncomptoUncomp method (operation 520).

[00146] **Fig. 6** is a flowchart illustrating operations 600 implemented by the ConnectUncomptoUncomp method. At operation 610 the topology loader 232 determines the upstream and downstream media types. As described above, in an exemplary implementation the media types may be retrieved from a data structure associated with the nodes. At operation 615 the topology loader finds the best match between the media types. By way of example, assume the upstream node indicates compatibility with three media types: IYUV, RGB32, and RGB16, while the downstream node indicates compatibility with RGB32 and YUYV media types. The best match is RGB32, so the topology loader would select RGB32 as the media type.



[00147] At operation 620 the topology loader 232 determines whether intermediate nodes such as, e.g., encoders or decoders, are required. If so, then control passes to operation 625 and the intermediate nodes are inserted and connected. Control then passes to operation 630, and the nodes are connected with the most compatible matching media type.

[00148] **Fig. 7** is a flowchart illustrating operations for handling optional nodes. Optional nodes are nodes which are added to the partial topology and marked as optional. If the topology loader 232 cannot resolve a connection to this node, then the node is skipped and the connection is created between the node previous to the optional node and the node behind it. The aim is that topology loading should not fail due to problems with the optional node.

[00149] In an exemplary implementation only transform nodes with one input and one output will be allowed to be set as optional. Other configurations would not allow the node to be skipped. A `SetOptionalFlag( BOOL bOptional , BOOL bIgnoreDownstreamType)` method may be added to the `IMFTopologyNode` for this. All the application needs to do is set this flag to true for the nodes it will consider optional. This may be done on the partial topology, either when it is created by the application, or if it isn't created by the application then it could access it thorough the topology callback.

[00150] Referring to Fig. 7, operations 700 for handling optional nodes may begin at the `ConnectNodes` module 710. When the `ConnectNodes` module is

invoked to connect a first (*i.e.*, upstream) node and a second (*i.e.*, downstream) node, control passes to operation 715, where it is determined whether the downstream node is optional. If the second (*i.e.*, downstream) node is optional, then control passes to operation 720, and the second node is placed on the optional node list. At operation 725 the analysis moves to the next downstream node (*i.e.*, the third node). Control then passes back to operation 710, and ConnectNodes is invoked with the first node and the third node.

[00151] The operations 710 through 725 are repeated until, at operation 715 the downstream node is not optional. In this event control passes to operation 730, where the optional node list created in operations 710 through 725 is empty. If the list is empty, then control passes to operation 745 and the nodes are connected using conventional connection logic. By contrast, if the list is not empty, then control passes to operation 735 where the non-optional nodes are connected. Control then passes to operation 740, and logic is executed to determine whether to connect the optional nodes between the non-optional nodes. In an exemplary implementation, an optional node will be included if its input type is the upstream preferred format, its output type is either the downstream preferred format, or the upstream preferred format. Further if there is more than one optional node, then the topology loader attempts to add the biggest group of optional nodes possible which meet these restrictions.

## **Exemplary Operating Environment**

[00152] The various components and functionality described herein are implemented with a number of individual computers. Fig. 8 shows components of a typical example of a computer environment 800, including a computer, referred to by reference numeral 802. The computer 802 may be the same as or different from computer 102 of Fig. 1. The components shown in Fig. 8 are only examples, and are not intended to suggest any limitation as to the scope of the functionality of the invention; the invention is not necessarily dependent on the features shown in Fig. 8.

[00153] Generally, various different general purpose or special purpose computing system configurations can be used. Examples of well known computing systems, environments, and/or configurations that may be suitable for use with the invention include, but are not limited to, personal computers, server computers, hand-held or laptop devices, multiprocessor systems, microprocessor-based systems, set top boxes, programmable consumer electronics, network PCs, network-ready devices, minicomputers, mainframe computers, distributed computing environments that include any of the above systems or devices, and the like.

[00154] The functionality of the computers is embodied in many cases by computer-executable instructions, such as software components, that are executed by the computers. Generally, software components include routines, programs,

objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Tasks might also be performed by remote processing devices that are linked through a communications network. In a distributed computing environment, software components may be located in both local and remote computer storage media.

[00155] The instructions and/or software components are stored at different times in the various computer-readable media that are either part of the computer or that can be read by the computer. Programs are typically distributed, for example, on floppy disks, CD-ROMs, DVD, or some form of communication media such as a modulated signal. From there, they are installed or loaded into the secondary memory of a computer. At execution, they are loaded at least partially into the computer's primary electronic memory.

[00156] For purposes of illustration, programs and other executable program components such as the operating system are illustrated herein as discrete blocks, although it is recognized that such programs and components reside at various times in different storage components of the computer, and are executed by the data processor(s) of the computer.

[00157] With reference to Fig. 8, the components of computer 802 may include, but are not limited to, a processing unit 804, a system memory 806, and a system bus 808 that couples various system components including the system memory to the processing unit 804. The system bus 808 may be any of several

types of bus structures including a memory bus or memory controller, a peripheral bus, and a local bus using any of a variety of bus architectures. By way of example, and not limitation, such architectures include Industry Standard Architecture (ISA) bus, Micro Channel Architecture (MCA) bus, Enhanced ISA (EISA) bus, Video Electronics Standards Association (VESA) local bus, and Peripheral Component Interconnect (PCI) bus also known as the Mezzanine bus.

[00158] Computer 802 typically includes a variety of computer-readable media. Computer-readable media can be any available media that can be accessed by computer 802 and includes both volatile and nonvolatile media, removable and non-removable media. By way of example, and not limitation, computer-readable media may comprise computer storage media and communication media. “Computer storage media” includes volatile and nonvolatile, removable and non-removable media implemented in any method or technology for storage of information such as computer-readable instructions, data structures, program modules, or other data. Computer storage media includes, but is not limited to, RAM, ROM, EEPROM, flash memory or other memory technology, CD-ROM, digital versatile disks (DVD) or other optical disk storage, magnetic cassettes, magnetic tape, magnetic disk storage or other magnetic storage devices, or any other medium which can be used to store the desired information and which can be accessed by computer 802. Communication media typically embodies computer-readable instructions, data structures, program modules or other data in a

modulated data signal such as a carrier wave or other transport mechanism and includes any information delivery media. The term “modulated data signal” means a signal that has one or more of its characteristics set or changed in such a manner as to encode information in the signal. By way of example, and not limitation, communication media includes wired media such as a wired network or direct-wired connection and wireless media such as acoustic, RF, infrared and other wireless media. Combinations of any of the above should also be included within the scope of computer readable media.

[00159] The system memory 806 includes computer storage media in the form of volatile and/or nonvolatile memory such as read only memory (ROM) 810 and random access memory (RAM) 812. A basic input/output system 814 (BIOS), containing the basic routines that help to transfer information between elements within computer 802, such as during start-up, is typically stored in ROM 810. RAM 812 typically contains data and/or software components that are immediately accessible to and/or presently being operated on by processing unit 804. By way of example, and not limitation, Fig. 8 illustrates operating system 816, application programs 818, software components 820, and program data 822.

[00160] The computer 802 may also include other removable/non-removable, volatile/nonvolatile computer storage media. By way of example only, Fig. 8 illustrates a hard disk drive 824 that reads from or writes to non-removable, nonvolatile magnetic media, a magnetic disk drive 826 that reads from or writes to

a removable, nonvolatile magnetic disk 828, and an optical disk drive 830 that reads from or writes to a removable, nonvolatile optical disk 832 such as a CD ROM or other optical media. Other removable/non-removable, volatile/nonvolatile computer storage media that can be used in the exemplary operating environment include, but are not limited to, magnetic tape cassettes, flash memory cards, digital versatile disks, digital video tape, solid state RAM, solid state ROM, and the like. The hard disk drive 824 is typically connected to the system bus 808 through a non-removable memory interface such as data media interface 834, and magnetic disk drive 826 and optical disk drive 830 are typically connected to the system bus 808 by a removable memory interface.

[00161] The drives and their associated computer storage media discussed above and illustrated in Fig. 8 provide storage of computer-readable instructions, data structures, software components, and other data for computer 802. In Fig. 8, for example, hard disk drive 824 is illustrated as storing operating system 816', application programs 818', software components 820', and program data 822'. Note that these components can either be the same as or different from operating system 816, application programs 818, software components 820, and program data 822. Operating system 816', application programs 818', software components 820', and program data 822' are given different numbers here to illustrate that, at a minimum, they are different copies. A user may enter commands and information into the computer 802 through input devices such as a keyboard 836, and pointing

device (not shown), commonly referred to as a mouse, trackball, or touch pad. Other input devices may include source peripheral devices (such as a microphone 838 or camera 840 which provide streaming data), joystick, game pad, satellite dish, scanner, or the like. These and other input devices are often connected to the processing unit 802 through an input/output (I/O) interface 842 that is coupled to the system bus, but may be connected by other interface and bus structures, such as a parallel port, game port, or a universal serial bus (USB). A monitor 844 or other type of display device is also connected to the system bus 808 via an interface, such as a video adapter 846. In addition to the monitor 844, computers may also include other peripheral rendering devices (e.g., speakers) and one or more printers which may be connected through the I/O interface 842.

[00162] The computer may operate in a networked environment using logical connections to one or more remote computers, such as a remote device 850. The remote device 850 may be a personal computer, a network-ready device, a server, a router, a network PC, a peer device or other common network node, and typically includes many or all of the elements described above relative to computer 802. The logical connections depicted in Fig. 8 include a local area network (LAN) 852 and a wide area network (WAN) 854. Although the WAN 854 shown in Fig. 8 is the Internet, the WAN 854 may also include other networks. Such networking environments are commonplace in offices, enterprise-wide computer networks, intranets, and the like.



[00163] When used in a LAN networking environment, the computer 802 is connected to the LAN 852 through a network interface or adapter 856. When used in a WAN networking environment, the computer 802 typically includes a modem 858 or other means for establishing communications over the Internet 854. The modem 858, which may be internal or external, may be connected to the system bus 808 via the I/O interface 842, or other appropriate mechanism. In a networked environment, program modules depicted relative to the computer 802, or portions thereof, may be stored in the remote device 850. By way of example, and not limitation, Fig. 8 illustrates remote software components 860 as residing on remote device 850. It will be appreciated that the network connections shown are exemplary and other means of establishing a communications link between the computers may be used.

## **Conclusion**

[00164] Although the described arrangements and procedures have been described in language specific to structural features and/or methodological operations, it is to be understood that the subject matter defined in the appended claims is not necessarily limited to the specific features or operations described. Rather, the specific features and operations are disclosed as preferred forms of implementing the claimed present subject matter.